

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Jean-Marc Roch  
Victoria E.S. Scott  
Kristi L. Anderson  
James P. Sullivan
- 10 (ii) TITLE OF INVENTION: HUMAN ENDOSULFINE GENE
- (iii) NUMBER OF SEQUENCES: 13
- 15 (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Abbott Laboratories  
(B) STREET: 100 Abbott Park Road  
(C) CITY: Abbott Park  
(D) STATE: Illinois  
(E) COUNTRY: USA  
20 (F) ZIP: 60064-3500
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: Macintosh  
25 (C) OPERATING SYSTEM: System 7.0.1  
(D) SOFTWARE: MS Word
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER:  
30 (B) FILING DATE:  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Priscilla E. Porembski  
35 (B) REGISTRATION NUMBER: 33,207  
(C) DOCKET NUMBER: 6024.US.01
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 708/937-0378  
40 (B) TELEFAX: 708/938-2623  
(C) TELEX:

## (2) INFORMATION FOR SEQ ID NO:1:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1093 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear
- 50 (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

55 GTGACAGGAG CCGAAGCAGC AGCGCAGGTT GTCCCCGTTT CCCCTCCCCC TTCCCTTCTC 60  
CGGTTGCCTT CCCGGGCCCC TTACTCTCCA CAGTCCCGGT CCCGCCATGT CCCAGAAACA 120  
AGAAGAAGAG AACCTGCGG AGGAGACCGG CGAGGAGAAG CAGGACACGC AGGAGAAAGA 180  
AGGTATTCTG CCTGAGAGAG CTGAAGAGGC AAAGCTAAAG GCCAAATACC CAAGCCTAGG 240  
ACAAAAGCCT GGAGGCTCCG ACTTCCTCAT GAAGAGACTC CAGAAAGGGC AAAAGTACTT 300  
TGACTCAGGA GACTACAACA TGGCCAAAGC CAAGATGAAG AATAAGCAGC TGCCAAGTGC 360  
AGGACCAGAC AAGAACCTGG TGA CTGGTGA TCACATCCCC ACCCCACAGG ATCTGCCCA 420  
60 GAGAAAGTCC TCGCTCGTCA CCAGCAAGCT TCGGGGTAA CCTGAGCCCC CCTCTCCTCC 480

CCTTCCTCAA CCACTGGACG TTTATATATT ATAGGCAGGG ATGAAATGGG CACCTAGTCA 540  
 GATCTTCTCA GCTTGCTAGC CAGAAATGAC TGTGATTCTG CTGGGGGCTG CTGAGAAGGT 600  
 AATGTAGGTT GAAAAGGGGC TCTAAGTTTA TTTATTTCTG TAGATTGACA CTTCCACACA 660  
 CTCCCTGTAG TCCAGGTAGG GCCTAGAAAAT AGGAAAGGCT AGGATTGGAT AATGCTGCAA 720  
 5 ATGCTTTTTT TGTGTGAGAA ACTGGAGAGA TGTGATTCTT CCTTTTGGGA GAGAATGTCC 780  
 CAAAATTGAT TAGGCTGAGC CTTGGGAATA GTTTGGCAGG TTAAACATCC CAAGGCTAAC 840  
 CTAACGTAGT TGGGAAAGGT AGATTGAATG AGACATGTTT TCTGTGCTTC TAAGTGTTCT 900  
 GTCCCTTAGG CTGCTATTGC TTCATGTTTC CATTATGGCA GGTTTAGAGA ATCCTTAAAA 960  
 AGAAAAATTG ACTTGCTTGC CTAAACTAC AGTGCCCCCT TAGCCTCCAT TACTTAGTAT 1020  
 10 CTCTTACAGT TTGCTCTGGC TCTCAAATAA TATAAAGATT GATGAACATT ATTCACAAAA 1080  
 AAAAAAAAAA AAAGGGCGGC GCG 1103

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1199 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTGACAGGAG CCGAAGCAGC AGCGCAGGTT GTCCCCGTTT CCCCTCCCCC TTCCCTTCTC 60  
 CGGTTGCCTT CCCGGGCCCC TTACTACTCCA CAGTCCCGGT CCCGCCATGT CCCAGAAACA 120  
 25 AGAAGAAGAG AACCTGCGG AGGAGACCGG CGAGGAGAAG CAGGACACGC AGGAGAAAGA 180  
 AGGTATTCTG CCTGAGAGAG CTGAAGAGGC AAAGCTAAAG GCCAAATACC CAAGCCTAGG 240  
 ACAAAGCCTT GGAGGCTCCG ACTTCCTCAT GAAGAGACTC CAGAAAGGGC AAAAGTACTT 300  
 TGACTCAGGA GACTACAACA TGGCCAAAGC CAAGATGAAG AATAAGCAGC TGCCAAGTGC 360  
 AGGACCAGAC AAGAACCTGG TGAAGGTGA TCACATCCCC ACCCCACAGG ATCTGCCCCA 420  
 30 GAGAAAGTCC TCGCTCGTCA CCAGCAAGCT TGGCGGTGGC CAAGTTGAAT GATGCTGCCC 480  
 GGGGCTCTGC CAGATCCTGA GACTGCTTTT CCGGCTTCCC CTCCCTGCCC CACCCGGGTC 540  
 CTGTGCTGGC TCCTGCCCCT TCAGCCAGGG GTCAGGAGGT GGCTCGGGTG TGGGCTGGAG 600  
 AGGCAGAAGC CCTTTCCTGT TGGTGTCCCA GCACATGGAG CCCCTTGGGC TGAGCACCAA 660  
 GACCTTGAAC CTTTTTTTGT TTACCTTTTT TCCAAATAAC AGTTGGGAGA AATATCAATG 720  
 35 AAATTCTGGG GGTGGGGGTG GGGCTGAAAG GGTGGGGTGG GAGATATGGA GGAGTATGAA 780  
 TTAGGGCTTG GAGTTGGTAA AAACATTCCCT GACTATCCTT CTTAACCACG TGGCTGATGT 840  
 GGGGTAGGTA TGAGGGGAAG GAAGTGGAGT AGCCTAATGA AAAGGGGTTC TAGTTGAGCT 900  
 CTGTAGATAA ATGCCTTGTT TCAGTGTGGT TGGAGACCTG GTGTCAGATA AAAGAACTC 960  
 CATCCGCACA GACAGATGCA AACAGTCCCT CTAGTTCTGC AGAGCTAGTT GAGAACTCAA 1020  
 40 CATTAAATCAT TTTAAAAAGT ACTGTCTTGT AAATAGATTT GCTGTGGGAA GAAGGGCAGT 1080  
 GAGTGTGGGA GAAAGGAGCC GTGAGCGTGG GGAACCCAC AGAGCCCAA GGACTTTTTTC 1140  
 AGTATTTCGAA ATAAACAAAA CAAAAACCCA TGAAAAAACC CAAAAAATAA AAAAAAATA 1199

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Ser Gln Lys Gln Glu Glu Asn Pro Ala Glu Glu Thr Gly Glu  
 1 5 10 15  
 Glu Lys Gln Asp Thr Gln Glu Lys Glu Gly Ile Leu Pro Glu Arg Ala  
 20 25 30  
 Glu Glu Ala Lys Leu Lys Ala Lys Tyr Pro Ser Leu Gly Gln Lys Pro  
 35 40 45  
 60 Gly Gly Ser Asp Phe Leu Met Lys Arg Leu Gln Lys Gly Gln Lys Tyr

50 55 60  
 Phe Asp Ser Gly Asp Tyr Asn Met Ala Lys Ala Lys Met Lys Asn Lys  
 65 70 75 80  
 Gln Leu Pro Ser Ala Gly Pro Asp Lys Asn Leu Val Thr Gly Asp His  
 85 90 95  
 Ile Pro Thr Pro Gln Asp Leu Pro Gln Arg Lys Ser Ser Leu Val Thr  
 100 105 110  
 Ser Lys Leu Ala Gly  
 115

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Gln Lys Gln Glu Glu Glu Asn Pro Ala Glu Glu Thr Gly Glu  
 1 5 10 15  
 Glu Lys Gln Asp Thr Gln Glu Lys Glu Gly Ile Leu Pro Glu Arg Ala  
 20 25 30  
 Glu Glu Ala Lys Leu Lys Ala Lys Tyr Pro Ser Leu Gly Gln Lys Pro  
 35 40 45  
 Gly Gly Ser Asp Phe Leu Met Lys Arg Leu Gln Lys Gly Gln Lys Tyr  
 50 55 60  
 Phe Asp Ser Gly Asp Tyr Asn Met Ala Lys Ala Lys Met Lys Asn Lys  
 65 70 75 80  
 Gln Leu Pro Ser Ala Gly Pro Asp Lys Asn Leu Val Thr Gly Asp His  
 85 90 95  
 Ile Pro Thr Pro Gln Asp Leu Pro Gln Arg Lys Ser Ser Leu Val Thr  
 100 105 110  
 Ser Lys Leu Ala Gly Gly Gln Val Glu  
 115 120

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAAGAGACTC CAGAAAGGGC

20

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAAGCAATAG CAGCCTAAGG

20

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGTGCTGGGA CACCAACAG

19

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGAAGGCGCC ATGTCCCAGA AACAAGAAG

29

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CAAAGTCGAC TTACCCCGCA AGCTTGCTG

29

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCAAGTCGAC TCATTCAACT TGGCCACCC

29

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 111 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Ser Ala Glu Val Pro Glu Ala Ala Ser Ala Glu Glu Gln Lys Glu

1 5 10 15  
 Met Glu Asp Lys Val Thr Pro Glu Lys Ala Glu Glu Ala Lys Leu Lys  
 20 25 30  
 Ala Arg Tyr Pro His Leu Gly Gln Lys Pro Gly Gly Ser Asp Phe Leu  
 35 40 45  
 Arg Lys Arg Leu Gln Lys Gly Gln Lys Tyr Phe Asp Ser Gly Asp Tyr  
 50 55 60  
 Asn Met Ala Lys Ala Lys Met Lys Asn Lys Gln Leu Pro Thr Ala Thr  
 65 70 75 80  
 Pro Asp Lys Thr Glu Val Thr Gly Asp His Ile Pro Thr Pro Gln Asp  
 85 90 95  
 Leu Pro Gln Arg Lys Pro Ser Leu Val Arg Ser Lys Leu Ala Gly  
 100 105 110

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Glu Gly Ile Leu Pro Glu Arg Ala Glu Glu Ala Lys Leu Lys Ala Lys  
 1 5 10 15  
 Tyr Pro Ser Leu Gly Gln Lys Pro Gly Gly Ser Asp Phe Leu Met Lys  
 20 25 30  
 Arg Leu Gln Lys Gly Gln Lys Tyr Phe Asp Ser Gly Asp Tyr Asn Met  
 35 40 45  
 Ala Lys Ala Lys Met Lys Asn Lys Gln Leu Pro Ser Ala Gly Pro Asp  
 50 55 60  
 Lys Asn Leu Val Thr Gly Asp His Ile Pro Thr Pro Gln  
 65 70 75

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Glu Gly Ile Leu Pro Glu Arg Ala Glu Glu Ala Lys Xaa Xaa Ala Lys  
 1 5 10 15  
 Tyr Pro Ser Leu Gly Gln Lys Pro Gly Gly Ser Asp Phe Leu Met Lys  
 20 25 30  
 Arg Xaa Xaa Xaa Xaa Gln Lys Tyr Phe Asp Ser Gly Asp Tyr Asn Met  
 35 40 45  
 Ala Lys Xaa Xaa Xaa Xaa Xaa Xaa Gln Leu Pro Ser Ala Gly Pro Asp  
 50 55 60  
 Lys Asn Leu Val Thr Gly Asp His Ile Pro Thr Pro Gln Asp Leu Pro  
 65 70 75 80  
 Gln Arg